

RESEARCH AND LIBRARY PERSPECTIVES IN WORKING WITH DIGITAL CULTURAL HERITAGE COLLECTIONS

Ditte Laursen & Katrine Gasser
History & Digital Archives: Production, content
and usage, November 9th, 2017



**DET KGL.
BIBLIOTEK**
Royal Danish Library

RESEARCH AND LIBRARY PERSPECTIVES IN WORKING WITH DIGITAL CULTURAL HERITAGE COLLECTIONS

COLLABORATIONS
COLLECTIONS AND ACCESS
FACEBOOK AS EXAMPLE



**DET KGL.
BIBLIOTEK**
Royal Danish Library

COLLABORATIONS



**DET KGL.
BIBLIOTEK**
Royal Danish Library

COLLABORATIONS

RESEARCH
PESPECTIVE



**DET KGL.
BIBLIOTEK**

Royal Danish Library



RAMUND: A Century of Radio and Music in Denmark. Music Genres, Radio Genres, and Mediatisation

■ RAMUND

- ▶ Home
- ▶ About the project
- ▶ Participants
- ▶ Results
- ▶ Literature
- ▶ Resources
- ▶ Collaborators and related projects
- ▶ Contact
- ▶ Music Sound Radio International Seminar



British Post Office engineers inspect Marconi's wireless telegraphy (radio) equipment in 1897
(Wikimedia commons)

A Century of Radio and Music

The RAMUND radio research project is funded by The Danish Council for Independent Research. The funds will be used for seminars, publications, and for realizing the following 11 subprojects:

- Music, Radio, and Liveness
- Danish National Radio from the Outside: Negotiations of Professionalism and Amateurism
- Music Charts and DR
- Radio Hosts as Music Presenters
- SILENCE! ... What you listen to, is who you are
- Broadcasting Musical Alterity
- Genre Culture, Segmentation and Formats
- Negotiating Programming Policies for Classical Music during the Hot and Cold War
- Music on Interbellum National Radio
- Contemporary Sports Radio Music – An Examination of Formats and Functions
- The Evil Twins of 'Mechanical Music' Revisited – DR and Phonographic Music before World War II

Future activities

Publication of a popular history of music and radio in Denmark, *Stil nu ind ...: Musikken og Danmarks Radio* (January 2018)

Publication of an anthology of research articles, *Tunes for All? Music on Danish Radio* (April 2018)

Publication of papers from the international seminar (June 2018)

Closing workshop (two days, January 2018)...

[More details](#)

Publications

Krogh, Mads & Morten Michelsen (2016): Music, Radio, and Mediatization. *Media, Culture and Society*.

Michelsen, Morten (2016): Here, There and in Between: Radio Spaces before the Second World War. In Johannes Brusila, Bruce Johnson and John Richardson (eds.): *Memory, Space, Sound*. London: Intellect Press....

[More details](#)

INDIVIDUALS



**DET KGL.
BIBLIOTEK**

Royal Danish Library

Forskningsportal

Forside | **Projekter** | Presse/medie | Priser | Forskere | Forskningsenheder | Aktiviteter | Publikationer | Om portalen

Militærarealer i et historisk perspektiv [Engelsk](#)

Svenningsen, Stig Roar (Projektdeltager)
Levin, Gregor, Institut for Miljøvidenskab, Aarhus Universitet, Danmark (Projektdeltager)

[Det Kgl. Bibliotek](#)

Dette projekt omhandler den militære arealanvendelses betydning for naturkvalitet og naturbeskyttelse i Danmark. Projektet tager udgangspunkt i de projektelementer som er beskrevet i projektforslaget 5.5.1 Militærarealer i et historisk perspektiv (Kanstrup m.fl. 2015).

Kanstrup, Niels, Konsortiet Faunaforst, Orbicon, og Dansk Jagtakademi. 2015. Projektkatalog - militærnatur : synergi mellem militær arealanvendelse og naturudvikling. Hjørring: Forsvarsministeriets Ejendomsstyrelse.

Status	Igangværende
Periode	11/08/2017 → 31/12/2017
Finansieringskilde	Anden offentlig støtte (Offentlig) Forsvarsministeriets Ejendomsstyrelse

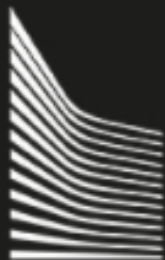
INDIVIDUALS



**DET KGL.
BIBLIOTEK**

Royal Danish Library

INDIVIDUALS



European
Commission

Horizon 2020
European Union funding
for Research & Innovation



**DET KGL.
BIBLIOTEK**
Royal Danish Library

CULTURAL HERITAGE CLUSTER



DeIC (Danish e-Infrastructure Cooperation) has been charged with spreading High-Performance Computing (HPC) to new research areas, such as the humanities and social science areas. In order to respond to this, DeIC and the Royal Danish Library have agreed to establish the **DeIC National Cultural Heritage Cluster, Royal Danish Library**.

The cultural heritage cluster applies state-of-the-art technologies within data science, and for the first time ever facilitates quantitative research projects on the digital Danish cultural heritage – e.g. radio and TV programmes, websites and historical newspapers.

In recent years, the Royal Danish Library has participated in national and international research infrastructure projects based on Danish digital cultural heritage. The library has knowledge and competences about what it takes to offer, for instance, data search for structures and patterns in large data sets.

The agreement between DeIC and the Royal Danish Library has a total financial framework of DKK 7.2 million over three years.

TECHNICAL



**DET KGL.
BIBLIOTEK**

Royal Danish Library

GOALS OF THE DIGITAL HUMANITIES LAB

- 1) Enhance and support digitally-enabled research across the humanities and arts
- 2) To develop, maintain and operate an infrastructure in support of ICT based research practices within the humanities and arts
- 3) To work with communities of practice to:



TECHNICAL

Knowledge, expertise, methodologies and practices across and related research infrastructures and institutions
- libraries, museums and archives in Denmark.



**DET KGL.
BIBLIOTEK**

Royal Danish Library

Tilsagn om deltagelse i projektet Trykkefriheden og en ny offentligheds tilblivelse



**DET KGL.
BIBLIOTEK**

RESEARCH & INNOVATION

Participant Portal

European Commission > Research & Innovation > Participant Portal > Opportunities

HOME FUNDING OPPORTUNITIES HOW TO PARTICIPATE PROJECTS & RESULTS EXPERTS SUPPORT

EU Programmes 2014-2020

Search Topics

TOPIC : European cultural heritage, access and analysis for a richer interpretation of the past.

COLLECTIONS



**DET KGL.
BIBLIOTEK**

Royal Danish Library

RESEARCH & INNOVATION
Participant Portal

Commission > Research & Innovation > Participant Portal > Opportunities

HOME FUNDING OPPORTUNITIES HOW TO PARTICIPATE PROJECTS & RESULTS EXPERTS SUPPORT

Programmes 2014-2020

Search Topics

Updates

Alerts

Horizon 2020

3rd Health Programme

Asylum, Migration and Integration Fund

Consumer Programme

CSME

European Statistics Programme

Hercule III Programme

Internal Security Fund - Borders

Internal Security Fund - Police

Justice Programme

Pilot Projects & Preparatory Actions

TOPIC: ERC Starting Grant

Topic identifier: ERC-2018-STG
Publication date: 24 July 2017

Types of action: ERC-STG Starting Grant
DeadlineModel: single-stage
Opening date: 03 August 2017 **Deadline:**

Horizon 2020

Topic Description

Scope:

Objectives

ERC Starting Grants are designed to support excellent Principal Investigators at the start of their own independent research team or programme. Applicants must demonstrate the ground-breaking nature, ambition and feasibility of their proposals.

Size of ERC Starting Grants

Starting Grants may be awarded up to a maximum of **EUR 1 500 000** for a period of up to 5 years. The maximum award is reduced pro rata temporis for projects of a shorter duration, or ongoing projects.

Living Archives
Research Project

MALMÖ UNIVERSITY

Home About Research Outputs People Collaborators

The project prototypes how digital archives can become resources for social and cultural change

15:30 No the ho how to

The Past's Future

The Past's Future

- Home
- About the project
- Researchers
- Collaboration
- Publications
- Negotiating Cultural Rights
- Contact

THE PAST'S FUTURE: DIGITAL TRANSFORMATIONS AND CULTURAL HERITAGE INSTITUTIONS

Digital access to a fundamental human right
Promoting individual freedom and empowerment and yielding important benefits, education is a fundamental human right and essential for the exercise of other human rights.

and educational resources are not only stored on paper. Cultural heritage is also being digitized, and digital transformations, and data editing and presenting them in new ways.

making culture accessible, for creating a new culture as the collective memory of a society in relation to methods and using knowledge.

Collaborating with three Danish national cultural heritage institutions – the National Gallery, the National Museum and the Royal Library – and with Professor Mette Schnapp, Director of metalab at Harvard University, we wish to analyze the impact of digital transformations on the framework of education in light of the digital transformations – and how these transformations affect the future of the past as collected, preserved and presented by cultural heritage institutions.

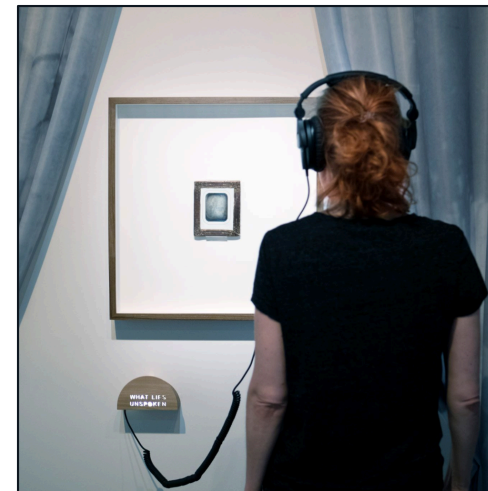
Internet Machine

Timo Arnall "Internet Machine" [elasticspace.com](#)

Uncertain Archives

Applying Cultural Theories of the Archive to Understand the Risks and Potential of Big Data.

data archives offer themselves as reassuring, neutral and innovative systems for the management of today's information deluge. Yet recent information scandals suggest that



ARCHIVES



COLLABORATIONS

RESEARCH
PESPECTIVE



**DET KGL.
BIBLIOTEK**

Royal Danish Library

COLLABORATIONS

ORGANISATIONAL
PESPECTIVE



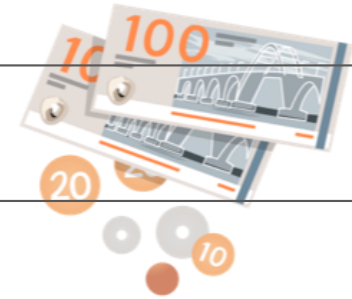
**DET KGL.
BIBLIOTEK**

Royal Danish Library

RESEARCH DEVELOPMENT



MORE RESEARCH



ONE INSTITUTION

information
specialist, curator,
librarian

it developer
technician

legal advisor

Researcher
Research Librarian

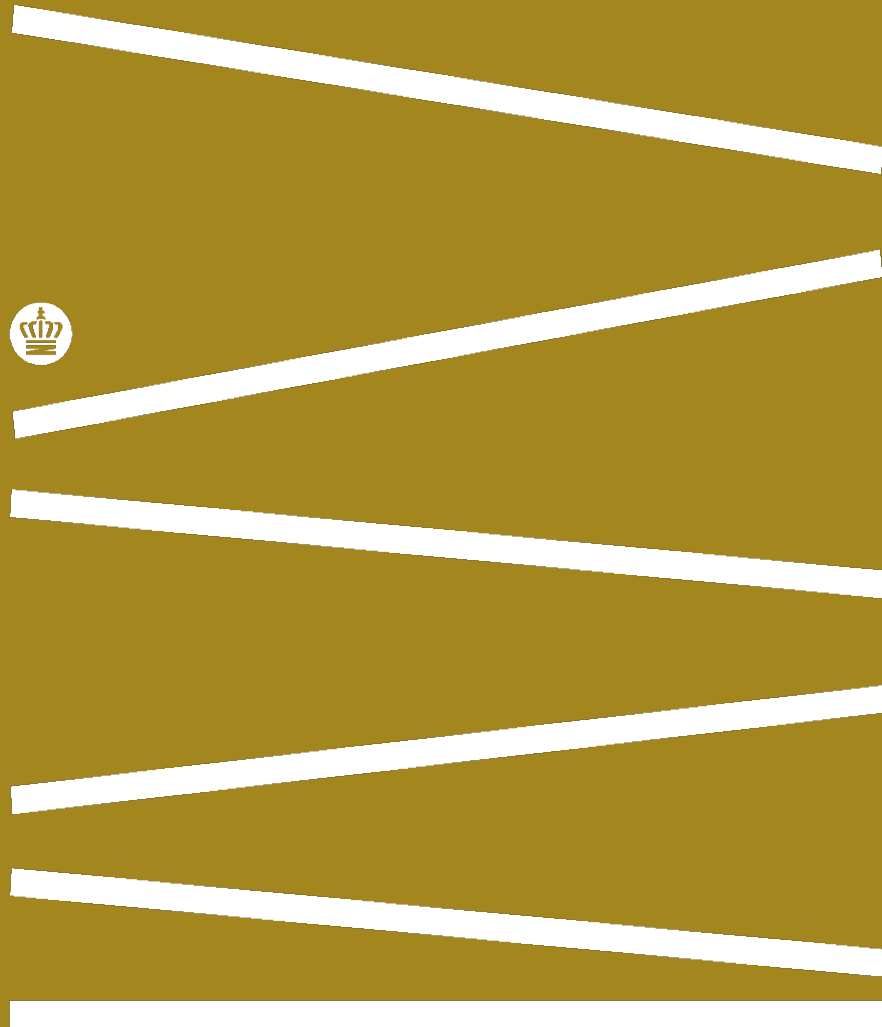


SYNERGIES LIBRARY BUSINESS



BETTER SERVICES





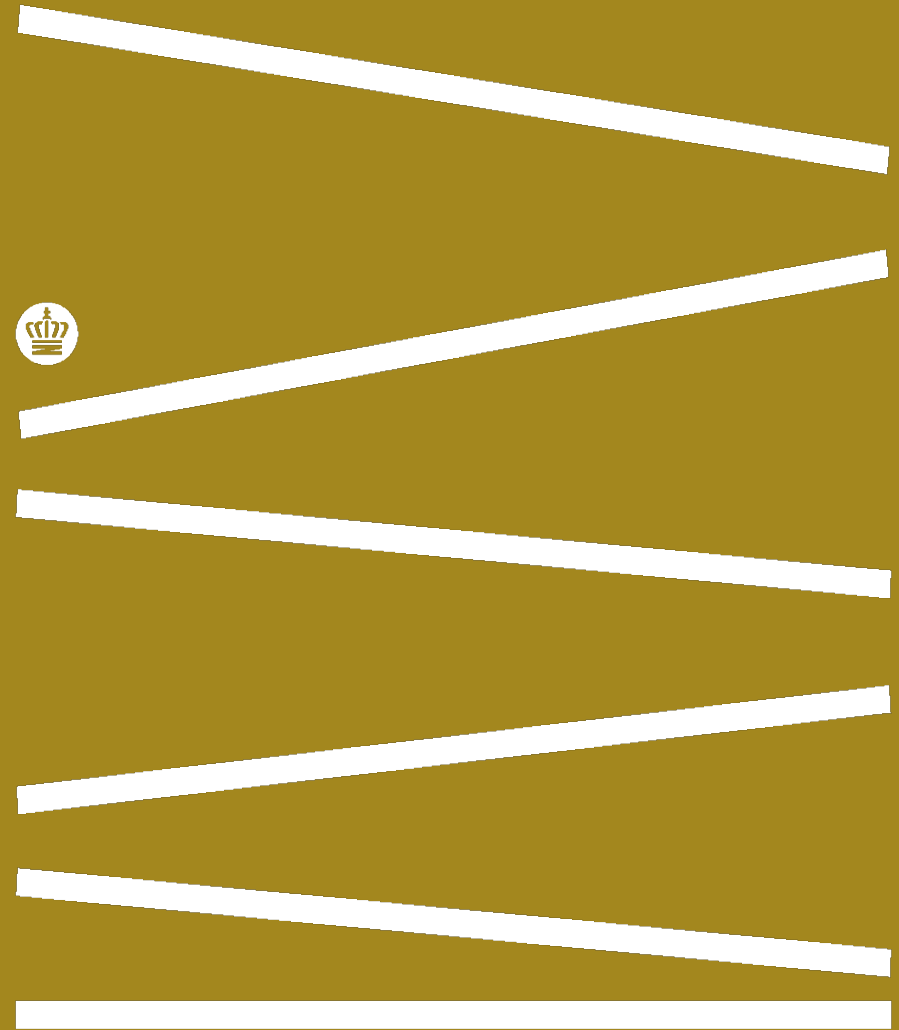
**DET KGL.
BIBLIOTEK**
Royal Danish Library

ACCESS COLLECTIONS



**DET KGL.
BIBLIOTEK**
Royal Danish Library

COLLECTIONS



**DET KGL.
BIBLIOTEK**
Royal Danish Library

DIGITAL BORN COLLECTIONS LEGAL DEPOSIT

The Danish webarchive

- 20+ billion "documents"
- 700 TB data
- Growth ~140 TB per year

Radio/TV

- 2.3 million hours of audio & video
- Over 2.5 PB of data
- Growth ~400 TB per year

Works

- books (40.000), audio books (10.000), newspapers (daily from 55 publishers), sheet music, texts (200) + pilot projects

Interactives

- Games (1100), Apps (30)

Audio

- all published CD's



DIGITAL BORN COLLECTIONS NON LEGAL DEPOSIT

MyArchive (20 people's emails)

Private archives



DIGITIZED COLLECTIONS

Photos

- 1.5 million fotos

Books

- 1000+ early prints
- 40.000 (digital on demand)

Journals

- 150+ journal titles

Manuscripts

- 1000+

Letters

- 40.000+

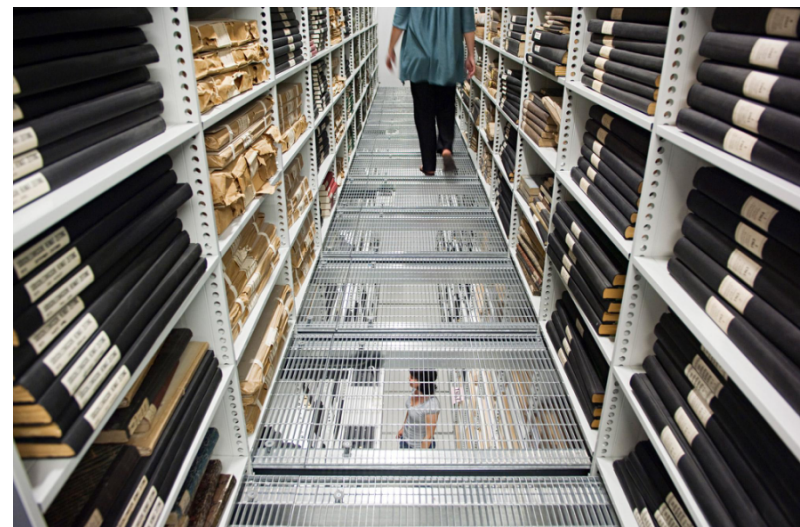
News Papers

- 35 million pages

Audio

- 50.000 sound files

etc.



ACCESS



**DET KGL.
BIBLIOTEK**
Royal Danish Library

[SEARCH ALL](#)
[NEWSPAPERS](#)
[RADIO](#)
[TV](#)
[COMMERCIALS](#)
[DANSK](#)

MEDI-STREAM

Search the Royal Danish Library's media collections

Search

Limit your search to collections you can [access](#).

HOW DO I FIND THE NEWSPAPERS?

You can now find newspapers in Mediestream. Click on 'Newspapers' in the menu. View a [list](#) of digitized newspapers. More newspapers are added during the coming years. Find information about [access](#) to the newspapers.

HOW DO I GET ACCESS?

You can search for information about each programme, newspaper page or commercial. At the Royal Danish Library - Copenhagen/Aarhus and the Danish Film Institute you can read, watch and listen to the contents. [Read more](#) about your options.

APPLICATIONS



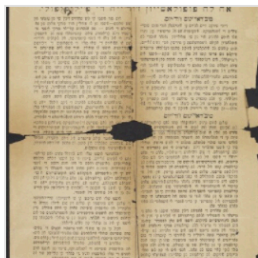
Fritekst

Search...

From yyyy-mm-dd

To yyyy-mm-dd

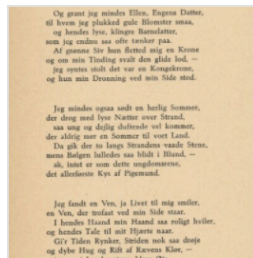
Search Q



Judaical Collection: Early and rare prints



H.C. Andersen



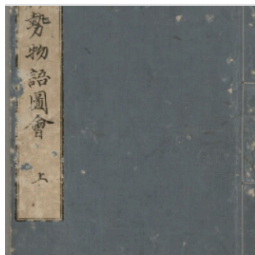
Pamphlets



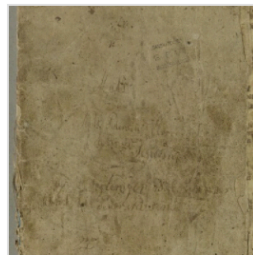
Political Manifestos



Maps and Atlases



Early prints from the Oriental Collection

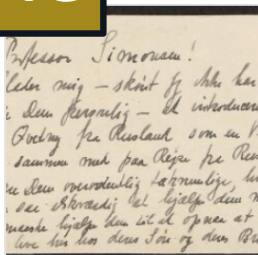


Danish Folklore Archives



Department Store Catalogues

APPLICATIONS



The Archive of David Simonsen



David Simonsens Manuscripts



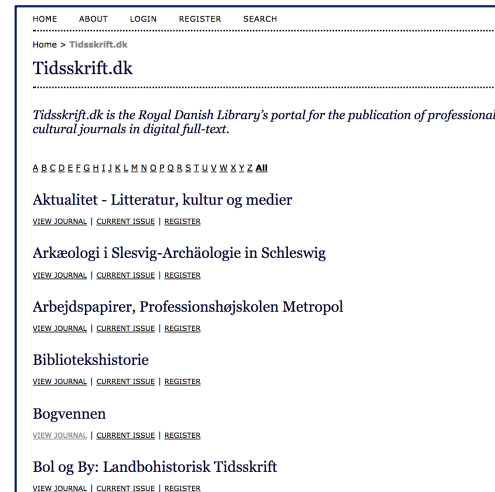
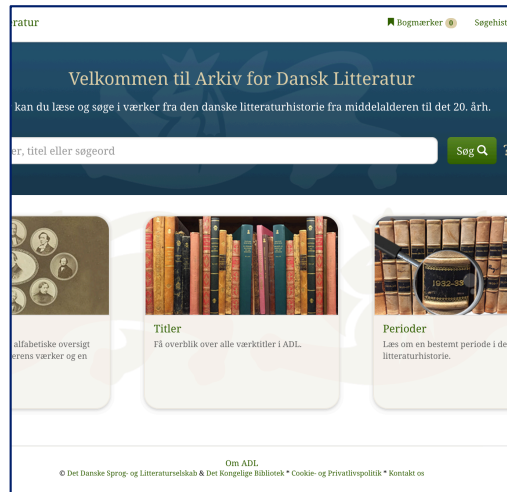
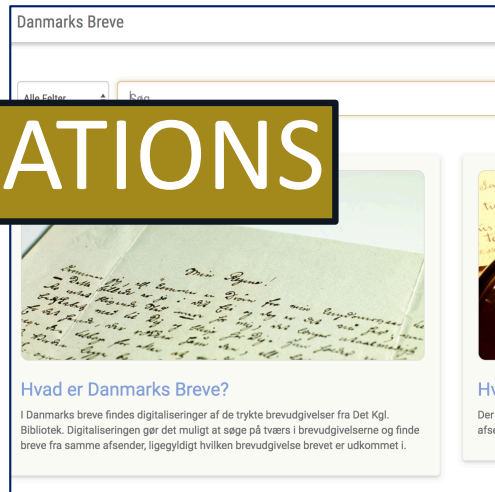
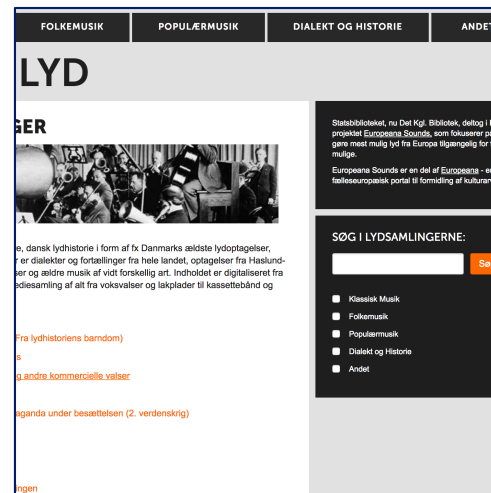
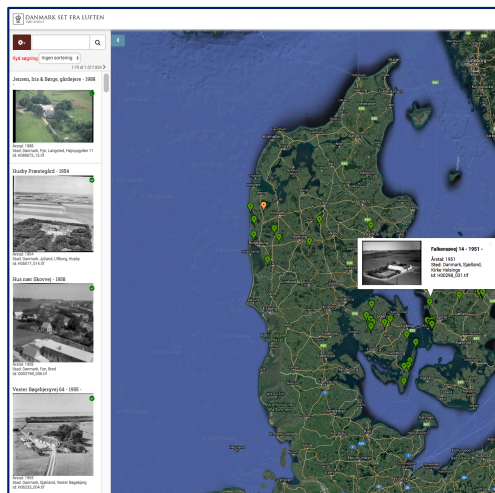
Images



Judaica Collections: Manuscripts



APPLICATIONS



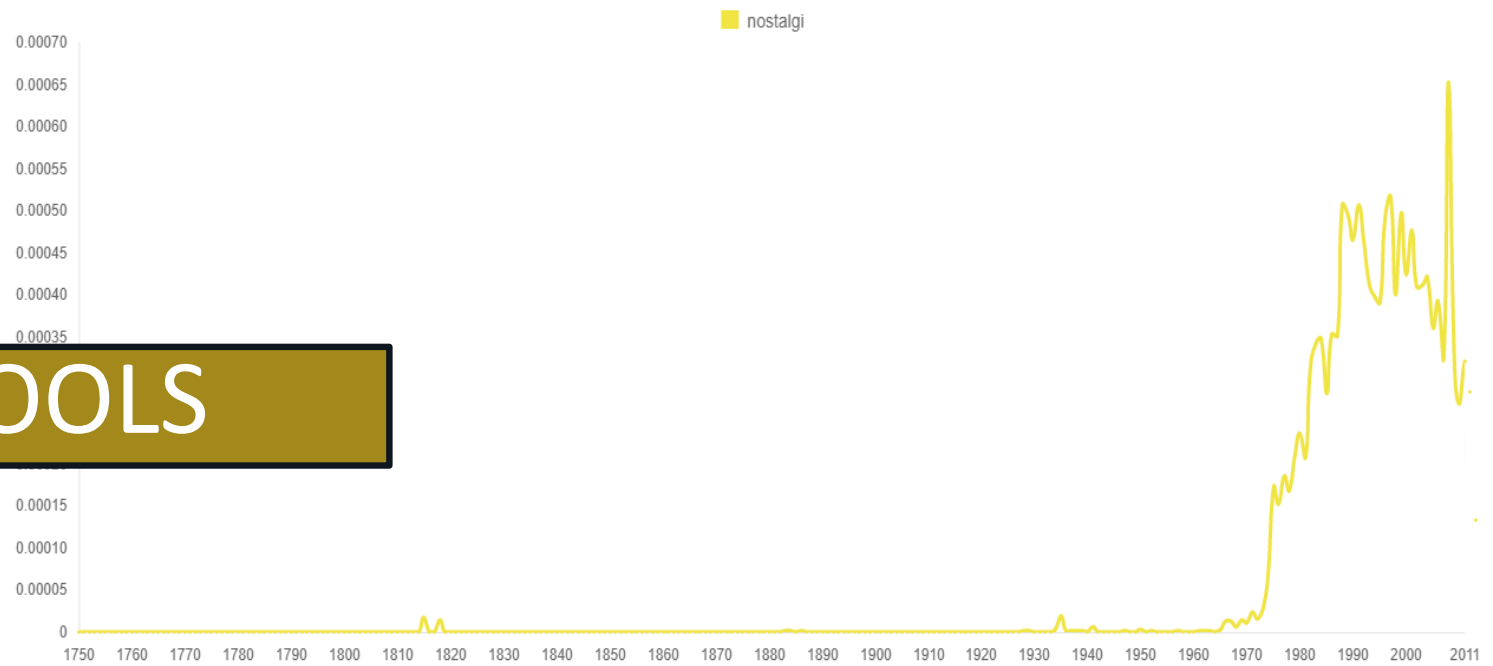
SMURF

nostalgi

Search

[Clear](#) [Download graph data](#)

No. of articles (%)



TOOLS



**DET KGL.
BIBLIOTEK**
Royal Danish Library

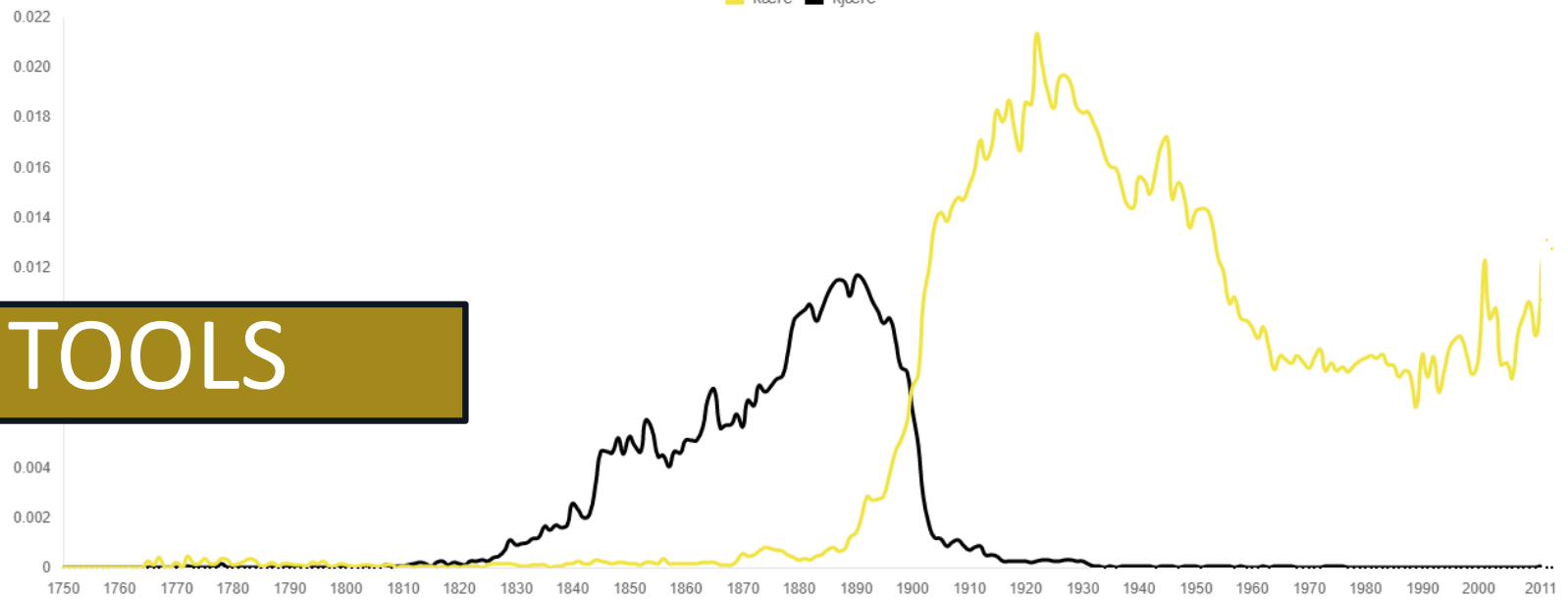
SMURF

kjære

Search

[Clear](#) [Download graph data](#)

No. of articles (%)



TOOLS



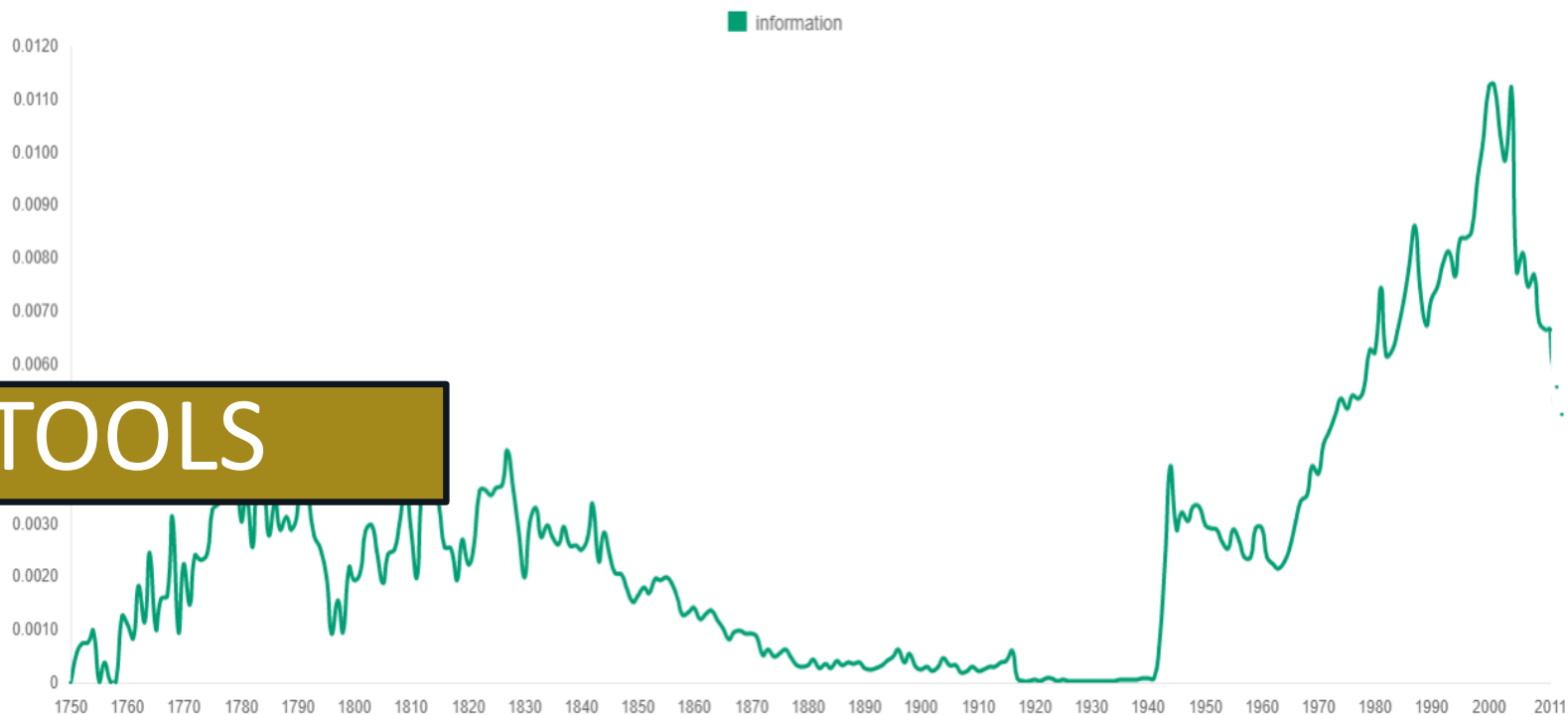
SMURF

information

Search

[Clear](#) [Download graph data](#)

No. of articles (%)



TOOLS

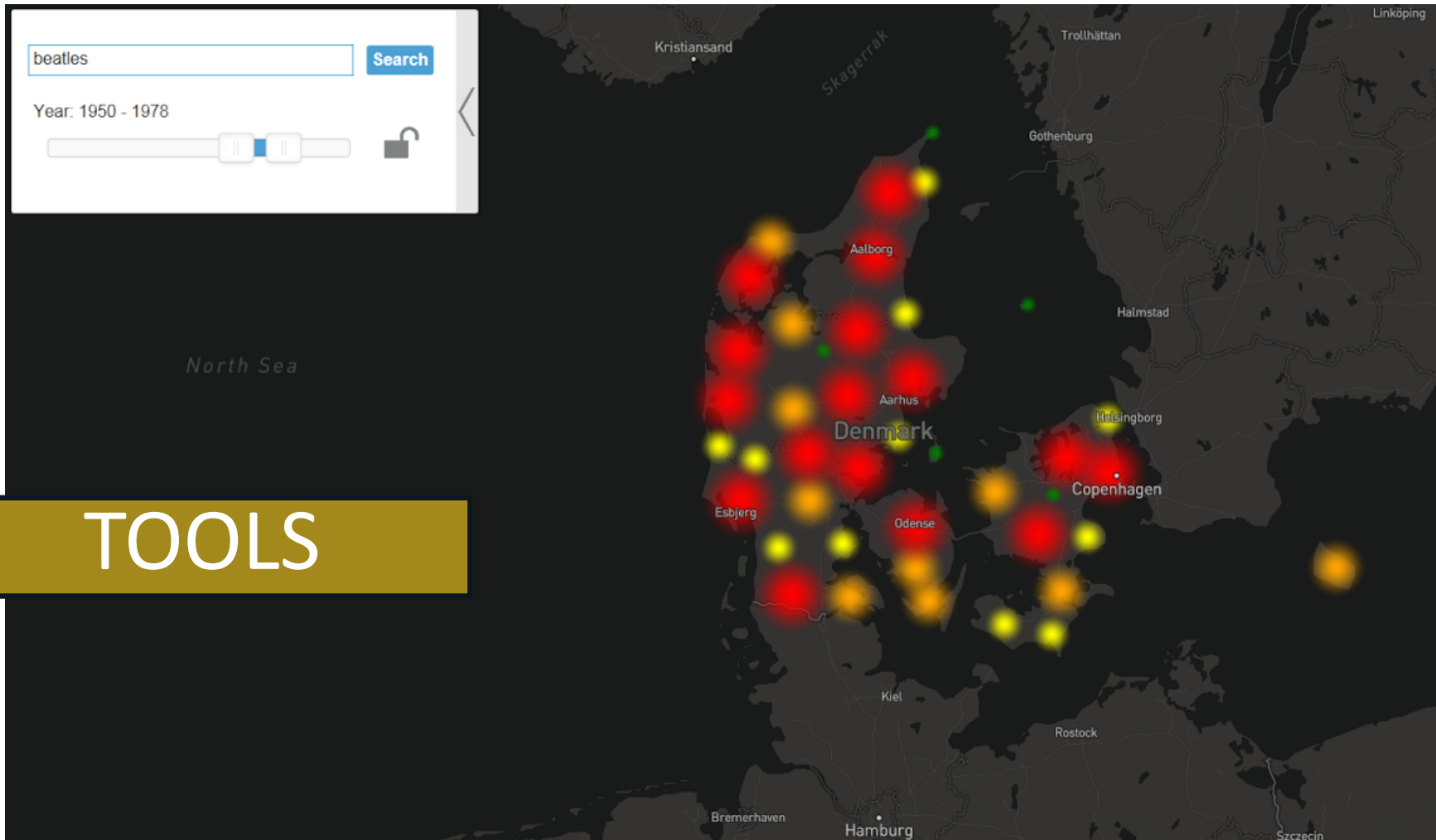


**DET KGL.
BIBLIOTEK**
Royal Danish Library

beatles

Year: 1950 - 1978

TOOLS



Newspapers from Royal Danish Library

Browse by

- [By Issue Date](#)
- [Authors](#)
- [Titles](#)
- [Subjects](#)

Search within this community and its collections: [GO](#)

A collection of OCR text based on newspapers digitised by the Royal Danish Library. The newspapers cover the years 1660 to 1877. The collection contains both newspapers printed in Denmark as well as newspapers printed in the former Danish-West Indies. The collection is sub-divided into years. Each year contains the newspapers which and which have been digitised. Please notice that the files covering the years 1679-1688 are missing.

Community

- [Newspapers 1749-1799](#)
Collection of OCR text in csv files from digitised newspapers
- [Newspapers 1800-1849](#)
Collection of OCR text in csv files from digitised newspapers
- [Newspapers 1850-1877](#)
Collection of OCR text in csv files from digitised newspapers

Recent Submissions

[Newspapers from 1877](#)

Royal Danish Library

NFWSPAPFRS FROM 1877 aalborgstiftstidende aarhusstiftstidende adresseavisen1759 aerneavis aktuelt

Search LOAR

[GO](#)

- Search LOAR
- This Community

[Advanced Search](#)

Browse

All of LOAR

[Communities & Collections](#)

[By Issue Date](#)

[Authors](#)

[Titles](#)

[Subjects](#)

This Community

[By Issue Date](#)

[Authors](#)

[Titles](#)

[Subjects](#)

My Account

[Login](#)

[Register](#)

Discover

Author

[Royal Danish Library \(81\)](#)

[Royal Danish Library \(61\)](#)

Subject

[berlingsketidende \(129\)](#)

[adresseavisen1759 \(119\)](#)

[aalborstiftstidende \(111\)](#)

DATA SETS



DET KGL. BIBLIOTEK DSFL COP ADL Documentation

ADL text search API ID and Relations fields

This document is a part of Royal Danish Library's APIs, and in particular The documentation on how use our texts. See also Licences & Legalese and Cavets

Search for
cat_ssi:work

result format

label	description
id	The ID of the record. It is th unless the record isn't refer as a string concatenation of sequence of arntids identify content indexed.
part_of_ssim	Array of IDs of trunk nodes hand. Typically containing <ul style="list-style-type: none"> • One (or more) work(s) contain works. • A volume as an ances

Filter fields

label	description
cat_ssi	Category of a text. Use whe omit otherwise
subcollection_ssi	Filter with respect to ADL s

APIs

DET KGL. BIBLIOTEK DSFL COP Documentation

Kort og Atlas Search (comma separate) Not before YYYY-MM-C Not aft

JSON	http://labs.kb.dk:8080/copapi/rest/api/cop?itemsPerPage=30&id=/maps/kortsa/2012/jul/kortatlas/da/&query=&after
RSS	http://www.kb.dk/cop/syndication/maps/kortsa/2012/jul/kortatlas/da?itemsPerPage=30&query=&page=1¬Befo
KML	http://www.kb.dk/cop/syndication/maps/kortsa/2012/jul/kortatlas/da?itemsPerPage=30&query=&page=1¬Befo
MODS	http://www.kb.dk/cop/syndication/maps/kortsa/2012/jul/kortatlas/da?itemsPerPage=30&query=&page=1¬Befo

Total number of items: 16866

Prev 1 2 3 4 5 ... 561 562 Next

Susatum ...; Warburgum elegans Westphalie Opp.

Circular Brunensis Pars Meridionalis

DET KGL. BIBLIOTEK DSFL COP Documentation

JSON	http://labs.kb.dk:8080/copapi/rest/api/dsfl?bbo=11.162109375,55.743347023896675,9.637756347656252,55.2
RSS	http://www.kb.dk/cop/syndication/images/luftfo/2011/maj/luftfoto/subject203?bbo=11.162109375,55.74334702
KML	http://www.kb.dk/cop/syndication/images/luftfo/2011/maj/luftfoto/subject203?bbo=11.162109375,55.74334702
MODS	http://www.kb.dk/cop/syndication/images/luftfo/2011/maj/luftfoto/subject203?bbo=11.162109375,55.74334702

These are 500 photos out of 179285



Individuel tailored access



FACEBOOK
AS
EXAMPLE



**DET KGL.
BIBLIOTEK**

Royal Danish Library

ARCHIVING TECHNIQUES

- still image
- screen filming
- harvesting via API
- html pages archived in the national web archive





Advantages

- look a like version
- robust format

Disadvantages

- cropped version
- not machine-readable
- not clickable
- no moving images
- no sound
- one moment in time



STILL IMAGE RQ's

- page layout can be identified and mapped out, including the potential culturally specific meanings that reside in the visual properties (cf. Pauwels, 2011, 2012)
- syntactical analysis of interrelated textual elements, mapping out weak or strong relations between elements and hierarchies between elements (cf. Brügger, 2007; Brügger, 2010b)
- the page's style (cf. Engholm, 2002) and how the page communicates through its graphical design, content design, interaction design and social design (cf. Engholm & Klastrup, 2010) ...





Advantages

- look a like version
- robust format
- moving images
- sound
- development over time
- timing of events

Disadvantages

- cropped version
- not machine-readable
- not clickable



SCREEN FILM RQ's

- research of movements and sounds, eg. all sorts of video and audio analysis of interactions and events
- moment-by-moment development of the web page and the timing of events
- micro-events in interacting with a facebook page, eg. how messages are edited before they are sent and equally how new messages come in while initial messages are still being typed (Meredith and Potter 2014)



The screenshot shows a web interface for 'Digital Footprints' with a sidebar on the left containing navigation options like 'Data', 'Participants', 'News Feeds', 'Profile Feeds', 'Groups', 'Feeds', 'Pages', 'Statistics', 'General', 'Content', 'Over Time', 'Force Directed Graph', 'Administration', 'Project Details', and 'Researchers'. The main area displays search results for 'må vi præsentere' on a Facebook page. It includes a search bar, a 'Toggle Comments' button, and a list of posts. The first post is from 'Voice - Danmarks største stemme (Tv show)' with the text 'Må vi præsentere DANMARKS STØRSTE STEMME. Send din hilsen til dejlige, skønne, dygtige & fabelagtige E TILLYKKE FRA HELE VOICE.' and a photo of a person. Below the photo is a 'Comments 920' button. The second post is from 'Danmarks største stemme (Tv show)' with the text 'Vi har netop fået fingrene i 14 ekstra billetter til lørdagens store semifinale på 'Voice'.' and a 'Comments 78' button.

Advantages

- capture of entire page's content
- capture of developments over time
- machine-readable (searchable, clickable, sortable)
- access to otherwise hidden data
- possible to manage data as big data

HARVESTING VIA API



The screenshot shows a Facebook interface for the page 'Voice - Danmarks største stemme'. The left sidebar contains navigation options like 'Data', 'Participants', 'News Feeds', 'Profile Feeds', 'Groups', 'Feeds', 'Pages', 'Statistics', 'General', 'Content', 'Over Time', 'Force Directed Graph', 'Administration', 'Project Details', and 'Researchers'. The main content area shows a post with a search bar containing 'må vi præsentere', a video thumbnail, and a comment section with 920 comments. Below the post, there is another comment from 'Danmarks største stemme' with 78 comments.

HARVESTING VIA API

Disadvantages

- no capture of the original visual display of elements
- videos are not playable (instead, link to live web)
- no capture of streamed content
- no capture of linked-to content (instead, forward to live web)
- proprietary format



API RESEARCH QUESTIONS

- classification: similarity of user profiles on different social networks (Malhotra et al. 2012)
- quantification: privacy awareness: low percentage of users change default privacy settings (Rauber et al. 2011)
- quantitative content analysis: Barack Obama portrayed more positively described across facebook groups than John McCain (Wolley et al. 2010)
- ...





Advantages

- capture of entire page
- machine-readable (searchable, clickable, sortable)
- access to otherwise hidden data
- capture of linked-to content (if included in the crawling)
- possible to manage data as big data
- robust format (warc/arc)
- look and feel like version (except for some applications, ie. flash, java script ...)





Disadvantages

- temporal delimitation
- spatial delimitation (to some extent)
- no capture of streamed content
- no capture of videos

WEB CRAWLING



WEB CRAWLING RQ'S

- All the kind of studies as with the previous methods, except for the video and audio analysis
- Link analysis: 22 % of webpages contain facebook URLs (Berk 2012)



OVERVIEW

	Replica version (look a like)	One version	Entire version (not cropped)	Native formats version	Temporal version
Screen shot image	x	x			
Screen shot pdf		x	x		
Screen movie (wall)	x	x			x
Screen movie (live chat)	x	x			x
Data extract from API (wall)		x	x	x	
Web crawler (wall)	x		x	x	
Web crawler (live chat)	x		x	(x)	



RESEARCHER

- Research question (iterative process)
- Knowledge about the collection
- Organisation
- Expertise (technical)
- Knowledge about access (legal)
- Possibilities

LIBRARY

- ← Marketing
- ← Dissimination
- ← Answers
- ← Advice
- ← Guidance



LINKS AND REFERENCES

- labs.statsbiblioteket.dk
- mediestream.dk
- <http://labs.kb.dk:8080/copapi/index.jsp>
- <http://en.statsbiblioteket.dk/kulturarvscluster>
- <http://digitalesamlinger.kb.dk/>
- <http://en.statsbiblioteket.dk/national-library-division/access-to-the-collections>
- <http://www.kb.dk/en/nb/materialer/e-ressourcer/index.html>

RESEARCH AND LIBRARY PERSPECTIVES IN WORKING WITH DIGITAL CULTURAL HERITAGE COLLECTIONS

Ditte Laursen & Katrine Gasser
History & Digital Archives: Production, content
and usage, November 9th, 2017



**DET KGL.
BIBLIOTEK**
Royal Danish Library